Plant Molecular Biology **27**: 249–261, 1995.

Cloning and developmental expression of pea ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit N-methyltransferase

Robert R. Klein [1] and Robert L. Houtz [2,*]
[1] *United States Department of Agriculture-Agriculture Research Service and* [2] *Department of Horticulture and Landscape Architecture, Plant Physiology/Biochemistry/Molecular Biology Program, University of Kentucky, Lexington, KY 40546, USA (*author for correspondence)*

Received 1 August 1994; accepted in revised form 26 September 1994

Key words: northern analysis, *Pisum sativum* L., protein methylation, protein methylase III (Rubisco LSMT, Rubisco large subunit N-methyltransferase, EC 2.1.1.43), Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco)

Abstract

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) large subunit (LS) N-methyltransferase (Protein methylase III, Rubisco LSMT, EC 2.1.1.43) catalyzes methylation of the ϵ-amino group of Lys-14 in the LS of Rubisco. With limited internal amino acid sequence information obtained from HPLC-purified peptic polypeptides from Rubisco LSMT, a full-length cDNA clone was isolated utilizing polymerase chain reaction-based technology and conventional bacteriophage library screening. The 1802 bp cDNA of Rubisco LSMT encodes a 489 amino acid polypeptide with a predicted molecular mass of ca. 55 kDa. A derived N-terminal amino acid sequence with features common to chloroplast transit peptides was identified. The deduced sequence of Rubisco LSMT did not exhibit regions of significant homology with other protein methyltransferases. Southern blot analysis of pea genomic DNA indicated a low gene copy number of Rubisco LSMT in pea. Northern analysis revealed a single mRNA species of about 1.8 kb encoding for Rubisco LSMT which was predominately located in leaf tissue. Illumination of etiolated pea seedlings showed that the accumulation of Rubisco LSMT mRNA is light-dependent. Maximum accumulation of Rubisco LSMT transcripts occurred during the initial phase of light-induced leaf development which preceded the maximum accumulation of *rbcS* and *rbcL* mRNA. Transcript levels of Rubisco LSMT in mature light-grown tissue were similar to transcript levels in etiolated tissues indicating that the light-dependent accumulation of Rubisco LSMT mRNA is transient. This is the first reported DNA and amino acid sequence for a protein methylase III enzyme.

Introduction

Ribulose-1,5-bisphosphate carboxylase-oxygenase (Rubisco) catalyzes the reduction of atmospheric CO_2 during photosynthesis. A great deal is known about the quaternary structure, catalytic mechanism, active site residues, *in vivo* regulatory mechanisms, and gene expression for this abun-

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number L34291 (rbcMT).

dant enzyme [for reviews see 1, 7, 29]. Higher-plant Rubisco is a hexadecameric protein composed of eight chloroplast-encoded LS and eight nuclear-encoded SS. Synthesis of the LS is accompanied by post-translational processing of the N-terminal domain [19, 30]. The N-terminal Met-1 and Ser-2 are removed and Pro-3 acetylated. Additionally, the LS of Rubisco from tobacco, muskmelon, pea, and several other species is post-translationally modified by trimethylation of the ϵ-amino group of Lys-14 [17, 19]. The enzyme responsible for this latter modification is a highly specific chloroplast-located *S*-adenosylmethionine (AdoMet):protein (lys) N-methyltransferase (Protein methylase III, Rubisco LSMT, EC 2.1.1.43). Recently, Rubisco LSMT was affinity-purified ca. 8000-fold from pea chloroplasts and identified as a monomeric protein with a molecular mass of ca. 57 kDa [44].

Protein methylation is a widespread and common post-translational modification catalyzed by several different protein methyltransferases [33]. Proteins which contain trimethyllysyl residues include cytochrome *c* [8–10], calmodulin [16, 38], histone [40, 42], and ribosomal proteins [5, 26]. However, the biological function of post-translational protein methylation in all but a few systems remains obscure. Trimethyllysine can serve as a metabolic precursor to carnitine [32], while carboxyl methylation of bacterial membrane proteins plays a major role in chemotaxis [6]. Evidence suggests that methylation of Lys-115 in calmodulin affects certain activities including *in vitro* NAD kinase activation [37], and *in vivo* susceptibilty to ubiquitination [14, 15 but see also 45, 46]. Conflicting reports [9, 11, 12 versus 3, 4] also implicate methylation of Lys-77 in cytochrome *c* as having a role in protein stability, heme incorporation, and mitochondrial transport. A major limitation to elucidating the biological role of lysine methylation in eukaryotes has been the abscence of a protein methylase III gene. Hence, molecular studies of the physiological and biochemical function performed by methylation of protein bound lysyl residues have been restricted to site-directed mutational analysis of the methylation site in the target protein [3, 4, 36]. These studies have been inconclusive as to the exact biological role of methylation of the ϵ-amine of protein bound lysyl residues.

The objectives of the present study were to isolate a cDNA of the Rubisco LSMT gene from pea and initiate studies of Rubisco LSMT gene expression. Utilizing amino acid sequence information derived from purified peptic polypeptide fragments from proteolyzed Rubisco LSMT, a full-length cDNA of Rubisco LSMT was obtained. The cDNA of Rubisco LSMT, rbcMT, was used to examine organ-specific and developmental parameters affecting rbcMT gene expression. The expression of two well characterized gene families, *rbcS* (SS of Rubisco) and *rbcL* (LS of Rubisco), were also examined to determine if rbcMT expression is coregulated with that of the Rubisco subunit genes, particularly the LS.

Materials and methods

Plant growth

Controlled environment-cultured peas (*Pisum sativum* L. cv. Laxton's Progress No. 9) were germinated and maintained in environmental chambers as described [44]. For developmental studies, seeds were either germinated at 23 °C in a dark chamber located in a light-tight room or were grown in an illuminated chamber with a light intensity of 300 μmol m^{-2} s^{-1} (incandescent plus fluorescent). After 8 days of growth in complete darkness, pea seedlings were either harvested into liquid nitrogen or were transferred to an illuminated chamber for a predetermined period prior to harvest.

Purification and assay of Rubisco LSMT

Rubisco LSMT was affinity-purified utilizing immobilized spinach Rubisco [44]. Briefly, purified spinach Rubisco [27] was immobilized to PVDF membranes (Millipore, Bedford, MA; 60 mg Rubisco/450 cm^2) which were then incubated for 4 h at 4 °C with pea chloroplast lysates (20 ml at 20 mg/ml protein per 450 cm^2 membrane). After incubation, the PVDF membranes were washed

with 50 mM Tris-HCl pH 8.2, 5 mM $MgCl_2$, 1 mM EDTA, 0.4 M NaCl and subsequently eluted with 20 ml of 50 mM Tris-HCl pH 8.2, 5 mM $MgCl_2$, 200 μM AdoMet and 50 μg/ml β-lactoglobulin per 450 cm^2 membrane. The eluent was concentrated by centrifugal ultrafiltration to a final volume of ca. 50 μl and used as a source for purified Rubisco LSMT. The yield from a single PVDF membrane containing immobilized spinach Rubisco was typically 7–10 μg of purified Rubisco LSMT. Assays of Rubisco LSMT activity were as previously described [18].

Peptide profiles and sequence analysis

To separate affinity-purified Rubisco LSMT from the β-lactoglobulin carrier protein, Rubisco LSMT was resolved by SDS-PAGE (10% acrylamide) prior to electrophoretic transfer to Immobilon-CD membranes (Millipore, Bedford, MA). Conditions for the electrophoretic transfer, visualization and subsequent *in vitro* enzymatic cleavage of Rubisco LSMT with pepsin were as described [34]. Peptic peptides released from Rubisco LSMT were isolated by reversed-phase HPLC as described [34] with an Aquapore RP-300 7 μm particle size octyl reversed-phase column (2.1 mm × 220 mm, Applied Biosystems, San Jose, CA). Peptic peptides were manually collected based on absorbance at 214 nm and samples reduced in volume to ca. 50 μl under vacuum. Amino acid sequence analyses were performed by the Macromolecular Structure Analysis Facility at the University of Kentucky, Lexington, using an Applied Biosystems 477A automated sequencer. For additional confirmation of amino acid sequence data, a duplicate sample of Rubisco LSMT was purified, proteolyzed, and peptic polypeptide fragments submitted for amino acid sequence analyses.

Synthesis of first-strand cDNA and polymerase chain reaction amplification

Pools of oligonucleotide primers encoding portions of two LSMT peptic peptides, P14 and P18, were synthesized with the number of different species (degeneracies) in each pool minimized as previously described [22]. First-strand cDNA synthesis and polymerase chain reaction (PCR) conditions were as described [22] except 5 μl of first-strand cDNA was used as PCR template and the PCR annealing temperature was reduced to 48 °C. The appropriate sense and antisense PCR primers directed against LSMT peptides, P14 and P18, are shown in Table 1. After amplification, the PCR product was purified and blunt-end ligated into SK plasmid (Stratagene, La Jolla, CA) and sequenced as described [24].

Screening of a pea cDNA library

To obtain a full-length cDNA of pea LSMT, a pea λgt10 cDNA library [13] was screened with the Rubisco LSMT-PCR product. Approximately 5×10^4 primary plaques were screened with a randomly labeled 360 bp PCR product of Rubisco LSMT under conditions recommended (Stratagene, La Jolla, CA). After four rounds of plaque purification, three potential positive plaques were identified. Following amplification and purification of bacteriophage DNA, Rubisco LSMT cDNAs were subcloned into SK plasmid and complete sequence of all three clones (ca. 1600 to 1775 bp in length) was obtained.

The technique of PCR-RACE (Rapid Amplification of cDNA Ends) was used to obtain a portion of the 5′region of LSMT essentially as described by the manufacturer (GIBCO-BRL, Gaithersburg, MD) except 100 ng of poly(A) mRNA was substituted for total RNA. The gene-specific (antisense) primer used to prime synthesis of first-strand LSMT cDNA was 5′-CCAAAAGAAGTCATCCAGCGTCAC (see Fig. 2, position 700–667). Amplification by PCR used the Anchor primer (supplied by GIBCO-BRL) and a second antisense LSMT-specific primer (5′-CAUCAUCAUCAUCCTGTGGC-AGAATACCAAAATAGT) which annealed to an internal, nested site within the LSMT cDNA (Fig. 2, position 515–492). The inclusion of the $(CAU)_4$ repeat sequence at the 5′ terminus per-

mitted a uracil DNA glycosylase (UDG) cloning strategy of the PCR-RACE product. PCR amplification conditions were as above except for an annealing temperature of 55 °C and an extension time of 40 s.

Northern blot analyses

Polyadenylated mRNA (0.5 μg per lane) or total RNA (2 μg per lane) was loaded on formaldehyde gels [39] and transferred to GeneScreen nylon membranes (DuPont-NEN, Wilmington, DE). Conditions for prehybridization and hybridization with radiolabeled antisense-RNA probes were as described [22]. The northern probe for *rbcS* was as described [22,24]; the probe for *rbcL* was as described [23]; the probe for rbcMT was a 1750 base antisense RNA from a portion of the open reading frame and 3′-untranslated region of pea.

Genomic Southern blot analyses

Nuclear DNA was isolated from nuclei as described [2]. Ten μg of high-molecular-weight DNA was digested to completion with *Eco* RI, *Hin*d III, and *Dra* I (50 units each). After digestion, DNA was ethanol-precipitated, electrophoresed on 0.8% agarose gels and transferred to Nytran nylon membranes using an alkaline transfer solution as described (Turboblotter instruction manual, Schleicher and Schuell, Keene, NH). Blots were prehybridized and hybridized at 42 °C in the presence of 50% formamide and 10% dextran sulfate. The probe was a random primer-labeled 1775 bp cDNA of pea LSMT (encompassing the open reading frame and entire 3′-untranslated region).

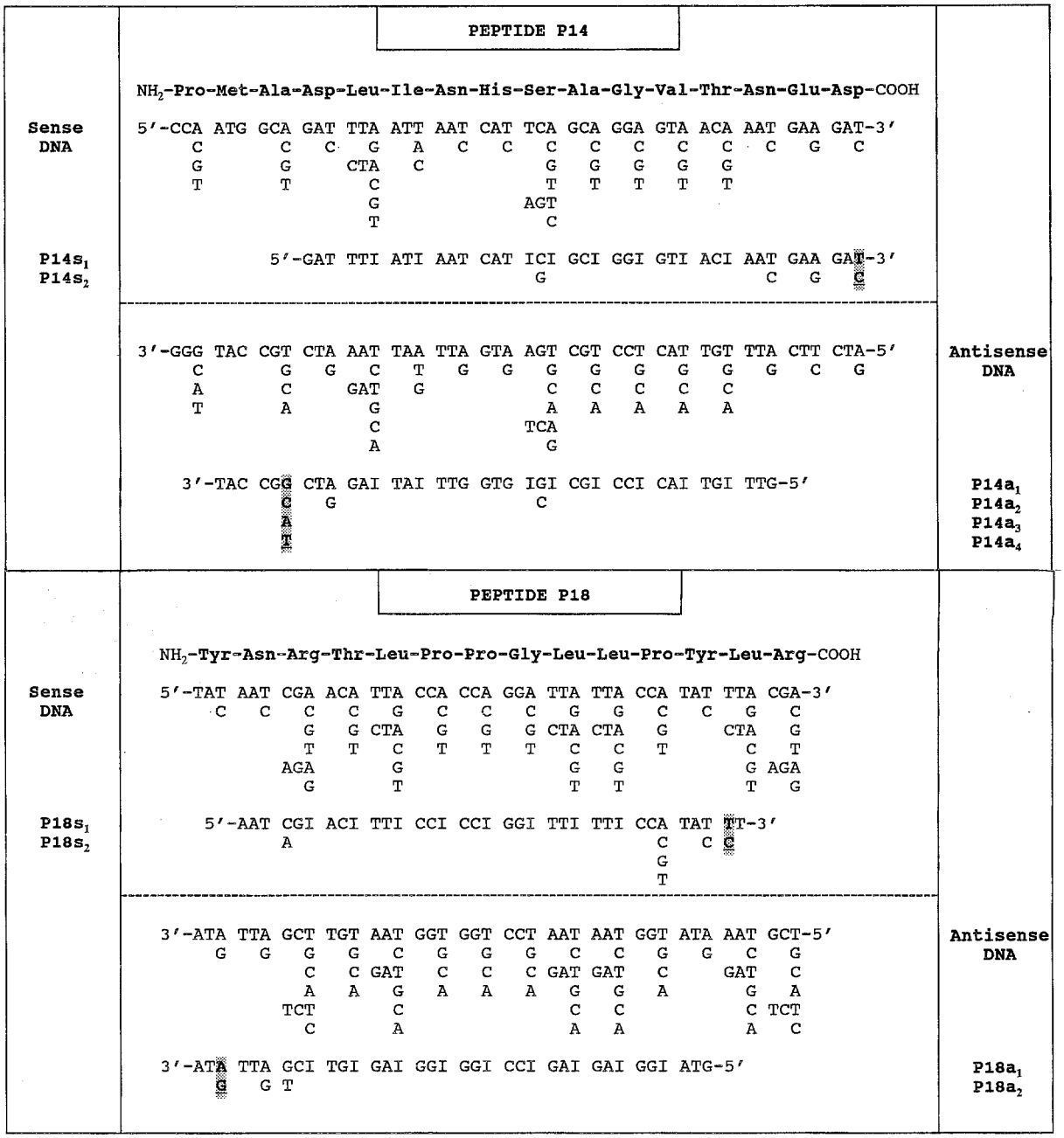
Miscellaneous

Computer alignment of the amino acid sequences was performed using the FastDB program (Intelligenetics, Mountain View, CA). Autoradiograms were scanned with an image acquisition densitometer (BioImage, Milligen/Biosearch, Ann Arbor, MI) to determine the relative intensity of mRNA signal and quantified on the basis of whole-band analysis.

Results

Rubisco LSMT has been purified ca. 8000-fold by a novel affinity purification technique from pea chloroplasts [44]. After affinity-purification of Rubisco LSMT, SDS-PAGE analysis showed a single polypeptide with an apparent molecular mass of ca. 57 kDa. Direct Edman degradative sequencing attempts followed by amino acid analyses after HCl hydrolysis of electroblotted affinity-purified Rubisco LSMT revealed that the N-terminus was blocked (data not shown). Thus, subsequent efforts were directed towards the acquisition of internal amino acid sequence as a starting point for isolating a cDNA of pea Rubisco LSMT. Reverse-phase HPLC isolation of peptic fragments from Rubisco LSMT resulted in the identification of several reliable amino acid sequences (Fig. 1, asterisks). One polypeptide peak, however, was heterogeneous and consisted of at least three subsequences which were identifiable based on differences in the relative amino acid yields after each cycle of sequencing. The partial amino-acid sequence of Rubisco LSMT enabled us to develop a molecular probe for the Rubisco LSMT gene (rbcMT) using PCR. Pools of deoxyinosine-containing primers encoding part of two peptic peptides, P14 and P18, were synthesized with the number of species in each pool minimized (Table 1). Using random-hexamer-primed first-strand cDNA as a template, the combination of primer pools P14–2s with P18–1a or P18–2a directed the synthesis of a single 360 bp PCR product (data not shown). No other primer combinations yielded a detectable PCR product. The fact that either antisense primer P18–1a or P18–2a (which differ by a single nucleotide near the 3′ terminus) directed the synthesis of a PCR product reflects the relative tolerance of the PCR system for base-pair mismatches near the 3′

Table 1. Degenerate PCR primer pools designed according to amino-acid sequence of Rubisco LSMT peptides P14 and P18. Underlined nucleotides represent degeneracy nearest 3′ terminus at which pools of primers differ.



terminus of the primer. The identity of the amplification product as a partial cDNA of rbcMT was confirmed by comparison of the deduced amino-acid sequence of the PCR product with additional peptic fragments from purified pea Rubisco LSMT protein (see Fig. 2).

The PCR-amplified fragment of rbcMT was used to screen a λgt10 pea cDNA library [13]. Three partial clones were obtained with inserts greater than 1600 bp in length. Complete sequence analysis of the three clones showed that the nucleotide sequence of all clones were identical. The

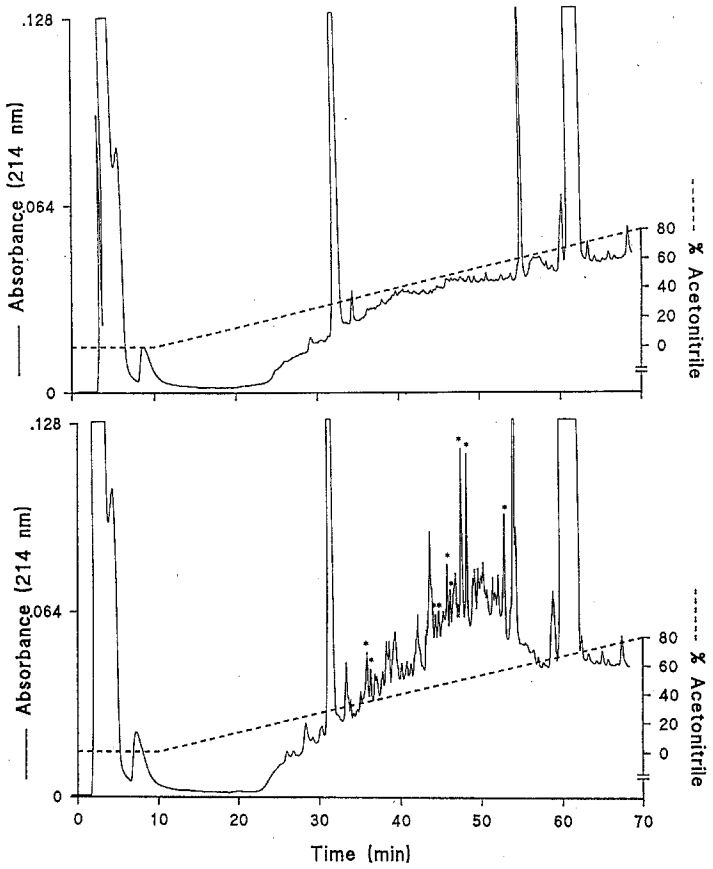
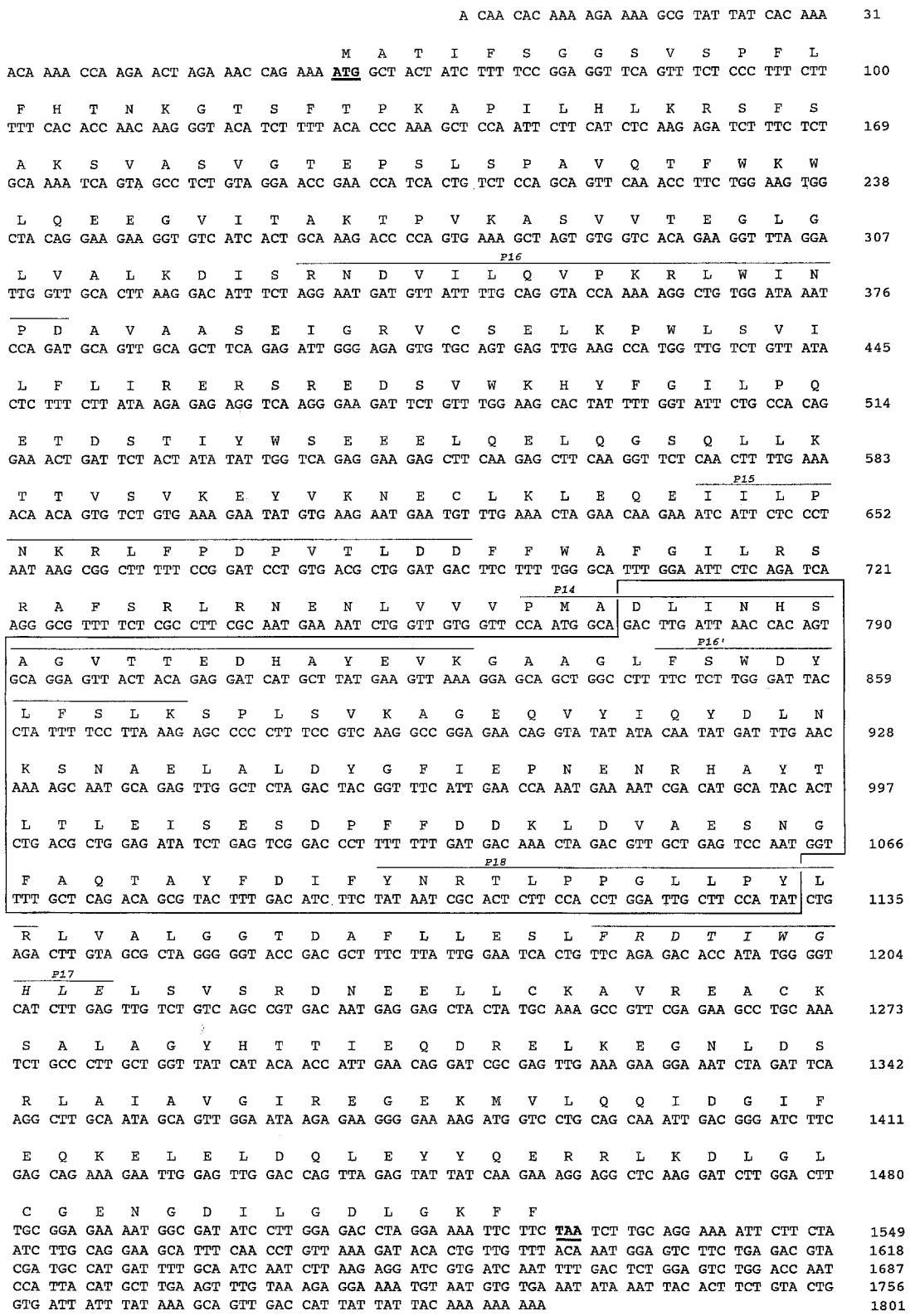


Fig. 1. Reversed-phase HPLC of peptic polypeptides from Rubisco LSMT. Top: control peptic digestion of Immobilon-CD membrane without Rubisco LSMT. Bottom: peptic digestion of affinity-purified Rubisco LSMT (ca. 30 μg) electroblotted to Immobilon-CD membrane as described in Materials and methods. Asterisks identify peaks with A_{214} absorbance which were collected and submitted for amino acid sequence analyses.

sequence of the PCR-derived cDNA was identical to the λgt10 cDNAs except for the incorrect identification of Thr-249 as an Asn residue during peptide sequencing of pepsin fragment P14. The longest clone (1775 bp in length) lacked only a portion of the 5′-untranslated region. The remainder of the 5′-untranslated region was obtained by PCR-RACE. The 515 bp PCR-RACE product was barely detectable on ethidium-stained gels which likely reflects the low abundance of the rbcMT mRNA in pea (data not shown). Sequence analysis confirmed the identity of the PCR-RACE product as encoding for the predicted 5′ portion of rbcMT including the remainder of the 5′-untranslated region. In the region where the PCR-RACE product overlapped the cloned cDNA of rbcMT, complete sequence identity was observed (Fig. 2, position 31–484 bp). Given these overlapping clones, we were able to assemble the sequence of the rbcMT cDNA as shown in Fig. 2. All of the peptic polypeptide sequences obtained from affinity-

Fig. 2. Nucleotide and predicted amino acid sequence of pea rbcMT cDNA. Nucleotide position is marked on the right. The start and stop codons are underlined and segments corresponding to peptic fragments are marked by lines above the amino-acid sequence. The position of amino acids encoded by the PCR-derived partial cDNA is blocked.


A CAA CAC AAA AGA AAA GCG TAT TAT CAC AAA 31
M A T I F S G G S V S P F L
ACA AAA CCA AGA ACT AGA AAC CAG AAA ATG GCT ACT ATC TTT TCC GGA GGT TCA GTT TCT CCC TTT CTT 100
F H T N K G T S F T P K A P I L H L K R S F S
TTT CAC ACC AAC AAG GGT ACA TCT TTT ACA CCC AAA GCT CCA ATT CTT CAT CTC AAG AGA TCT TTC TCT 169
A K S V A S V G T E P S L S P A V Q T F W K W
GCA AAA TCA GTA GCC TCT GTA GGA ACC GAA CCA TCA CTG TCT CCA GCA GTT CAA ACC TTC TGG AAG TGG 238
L Q E E G V I T A K T P V K A S V V T E G L G
CTA CAG GAA GAA GGT GTC ATC ACT GCA AAG ACC CCA GTG AAA GCT AGT GTG GTC ACA GAA GGT TTA GGA 307
P16
L V A L K D I S R N D V I L Q V P K R L W I N
TTG GTT GCA CTT AAG GAC ATT TCT AGG AAT GAT GTT ATT TTG CAG GTA CCA AAA AGG CTG TGG ATA AAT 376
P D A V A A S E I G R V C S E L K P W L S V I
CCA GAT GCA GTT GCA GCT TCA GAG ATT GGG AGA GTG TGC AGT GAG TTG AAG CCA TGG TTG TCT GTT ATA 445
L F L I R E R S R E D S V W K H Y F G I L P Q
CTC TTT CTT ATA AGA GAG AGG TCA AGG GAA GAT TCT GTT TGG AAG CAC TAT TTT GGT ATT CTG CCA CAG 514
E T D S T I Y W S E E E L Q E L Q G S Q L L K
GAA ACT GAT TCT ACT ATA TAT TGG TCA GAG GAA GAG CTT CAA GAG CTT CAA GGT TCT CAA CTT TTG AAA 583
P15
T T V S V K E Y V K N E C L K L E Q E I I L P
ACA ACA GTG TCT GTG AAA GAA TAT GTG AAG AAT GAA TGT TTG AAA CTA GAA CAA GAA ATC ATT CTC CCT 652
N K R L F P D P V T L D D F F W A F G I L R S
AAT AAG CGG CTT TTT CCG GAT CCT GTG ACG CTG GAT GAC TTC TTT TGG GCA TTT GGA ATT CTC AGA TCA 721
P14
R A F S R L R N E N L V V V P M A D L I N H S
AGG GCG TTT TCT CGC CTT CGC AAT GAA AAT CTG GTT GTG GTT CCA ATG GCA GAC TTG ATT AAC CAC AGT 790
P16'
A G V T T E D H A Y E V K G A A G L F S W D Y
GCA GGA GTT ACT ACA GAG GAT CAT GCT TAT GAA GTT AAA GGA GCA GCT GGC CTT TTC TCT TGG GAT TAC 859
L F S L K S P L S V K A G E Q V Y I Q Y D L N
CTA TTT TCC TTA AAG AGC CCC CTT TCC GTC AAG GCC GGA GAA CAG GTA TAT ATA CAA TAT GAT TTG AAC 928
K S N A E L A L D Y G F I E P N E N R H A Y T
AAA AGC AAT GCA GAG TTG GCT CTA GAC TAC GGT TTC ATT GAA CCA AAT GAA AAT CGA CAT GCA TAC ACT 997
L T L E I S E S D P F F D D K L D V A E S N G
CTG ACG CTG GAG ATA TCT GAG TCG GAC CCT TTT TTT GAT GAC AAA CTA GAC GTT GCT GAG TCC AAT GGT 1066
P18
F A Q T A Y F D I F Y N R T L P P G L L P Y L
TTT GCT CAG ACA GCG TAC TTT GAC ATC TTC TAT AAT CGC ACT CTT CCA CCT GGA TTG CTT CCA TAT CTG 1135
R L V A L G G T D A F L L E S L F R D T I W G
AGA CTT GTA GCG CTA GGG GGT ACC GAC GCT TTC TTA TTG GAA TCA CTG TTC AGA GAC ACC ATA TGG GGT 1204
P17
H L E L S V S R D N E E L L C K A V R E A C K
CAT CTT GAG TTG TCT GTC AGC CGT GAC AAT GAG GAG CTA CTA TGC AAA GCC GTT CGA GAA GCC TGC AAA 1273
S A L A G Y H T T I E Q D R E L K E G N L D S
TCT GCC CTT GCT GGT TAT CAT ACA ACC ATT GAA CAG GAT CGC GAG TTG AAA GAA GGA AAT CTA GAT TCA 1342
R L A I A V G I R E G E K M V L Q Q I D G I F
AGG CTT GCA ATA GCA GTT GGA ATA AGA GAA GGG GAA AAG ATG GTC CTG CAG CAA ATT GAC GGG ATC TTC 1411
E Q K E L E L D Q L E Y Y Q E R R L K D L G L
GAG CAG AAA GAA TTG GAG TTG GAC CAG TTA GAG TAT TAT CAA GAA AGG AGG CTC AAG GAT CTT GGA CTT 1480
C G E N G D I L G D L G K F F
TGC GGA GAA AAT GGC GAT ATC CTT GGA GAC CTA GGA AAA TTC TTC TAA TCT TGC AGG AAA ATT CTT CTA 1549
ATC TTG CAG GAA GCA TTT CAA CCT GTT AAA GAT ACA CTG TTG TTT ACA AAT GGA GTC TTC TGA GAC GTA 1618
CGA TGC CAT GAT TTT GCA ATC AAT CTT AAG AGG ATC GTG ATC AAT TTT GAC TCT GGA GTC TGG ACC AAT 1687
CCA TTA CAT GCT TGA AGT TTG TAA AGA GGA AAA TGT AAT GTG TGA AAT ATA AAT TAC ACT TCT GTA CTG 1756
GTG ATT ATT TAT AAA GCA GTT GAC CAT TAT TAT TAC AAA AAA AAA 1801

purified Rubisco LSMT were identified in the translated open-reading frame of the rbcMT cDNA. Independent confirmation of the identity of the cDNA has recently been obtained by cloning the complete open reading frame into an expression vector and synthesizing the polypeptide in *Escherichia coli* (P.G. Klein and R.L. Houtz, unpublished results). The recombinant polypeptide methylated spinach Rubisco *in vitro*, confirming the identity of the cDNA as encoding for Rubisco LSMT.

The rbcMT cDNA of 1802 bp in length contained a 5′ leader of 58 nucleotides which contained several short repeat elements and a 3′-untranslated region of 276 nucleotides. The rbcMT cDNA encoded for a protein of 489 amino acid residues with a predicted molecular mass of 55 kDa. Examination of the amino terminus of Rubisco LSMT revealed several motifs that commonly appear in chloroplast transit-peptide sequences, such as an abundance of hydroxylated amino acids Ser and Thr, presence of small hydrophobic amino acids, and general lack of acidic amino acids [21, 41]. Given that N-terminal sequence information could not be obtained for Rubisco LSMT, and that there is as yet no amino acid concensus sequence or secondary structural motif which unambiguously identifies the processing site for removal of chloroplast transit sequences [43], we were unable to determine the cleavage site between the precursor and mature forms of Rubisco LSMT.

Comparison of the deduced amino acid sequence of rbcMT cDNA with protein carboxyl methyltransferases from wheat (*D*-aspartyl/*L*-isoaspartyl protein methyltransferase, [28]) and *E. coli* (γ-glutamyl carboxyl methyltransferase, [31]) showed a low alignment score with sequence identity on the order of 10% (gaps in the sequence were introduced to maximize alignment). Three short amino acid regions (8 to 10 residues) of sequence similarity have been reported for several protein and small-molecule AdoMet-dependent methyltransferases [20]. Using manual alignment, we were unable to detect any of the three proposed sequence motifs of AdoMet-dependent methyltransferases in Rubisco rbcMT.

In a search of the SwissProt and NBRF-PIR data banks, the best match for Rubisco rbcMT was AfsR protein of *Streptomyces coelicolor* which reflected a 23% sequence identity over the entire protein, again with considerable gaps introduced.

DNA analysis

To obtain information on gene copy number, total pea leaf DNA was isolated and digested with several different restriction endonucleases (Fig. 3). A 1775 bp rbcMT cDNA probe hybridized to two *Eco* RI DNA fragments, ca. 5.3 kb and 2.0 kb (one *Eco* RI restriction endonuclease site is located within the sequenced cDNA). Two bands, ca. 3.5 kb and 1.3 kb, were observed after cleavage with *Dra* I, while a single band of 3.7 kb was observed after DNA digestion with *Hind* III.

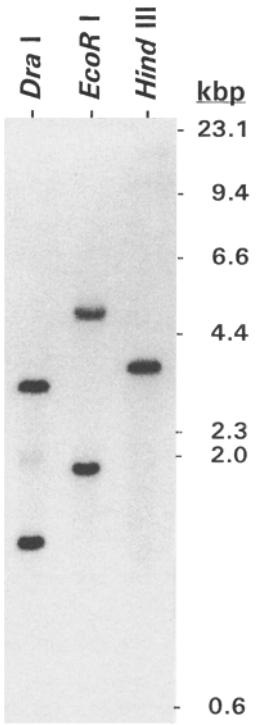


Fig. 3. Southern blot analysis of the rbcMT gene in pea. Ten μg of genomic DNA from pea was digested with *Eco* RI, *Hind* III, or *Dra* I, and electrophoresed on a 0.8% agarose gel. The blot was probed with a 1775 bp rbcMT cDNA of pea. Approximate sizes in kbp are indicated to the right. Blots were exposed to X-ray film for 48 h.

The simplicity of the DNA restriction digest pattern suggests that the gene copy number per haploid genome is low for rbcMT.

RNA analyses

Northern blot analyses were conducted on pea tissues to examine several developmental and organ-specific parameters governing rbcMT gene expression. As a basis for comparison, the expression of genes encoding Rubisco small (*rbcS*) and large (*rbcL*) subunit were concomitantly examined. The *rbcS* gene family and the *rbcL* gene were examined in an attempt to determine whether the expression of the Rubisco subunits and Rubisco LSMT was coordinated. Northern blot analysis indicated that the rbcMT gene encoded for a single species of mRNA of ca. 1.8 kb in length (see Fig. 4). Examination of organ-specific expression showed that accumulation of the rbcMT transcript paralleled the accumulation of *rbcL* and *rbcS* mRNA with the greatest proportion of mRNA being located in green leaf tissue. Transcripts encoding *rbcS*, *rbcL* and rbcMT were detected in pea stems, though the level of expression was 7, 10, and 28-fold lower, respectively, than in green leaves. The quantity of rbcMT, *rbcS*, and *rbcL* mRNA in root tissue was below the level of northern blot sensitivity. Maximum extractable Rubisco LSMT activity generally paralleled the accumulation of rbcMT mRNA though the enzyme activity detected in stems was greater than would be predicted based on mRNA levels. Maximum extractable Rubisco LSMT activity of roots, stems, and green leaves was 2, 15, and 36 pmole CH_3 per minute per mg protein, respectively. Finally, it should be noted that the exposure times of the rbcMT, *rbcS*, and *rbcL* northern analyses differ considerably and hence should be considered when comparing the absolute amounts of each transcript. The exposure time of rbcMT northern analyses were consistently 25 to 50 times longer than that of *rbcL* or *rbcS* suggesting that rbcMT transcripts do not accumulate to the level of the Rubisco subunits.

Examination of the accumulation of rbcMT mRNA during the greening of pea leaves is shown in Fig. 5. A low level of rbcMT mRNA was detected in 8-day-old dark-grown pea leaves (lane 1). Upon illumination of etiolated peas, rbcMT transcript levels increased ca. 3-fold after 24 h of illumination and then declined slightly after an additional 48 h of development in the light (lanes 2 and 3). The maximum extractable activity of

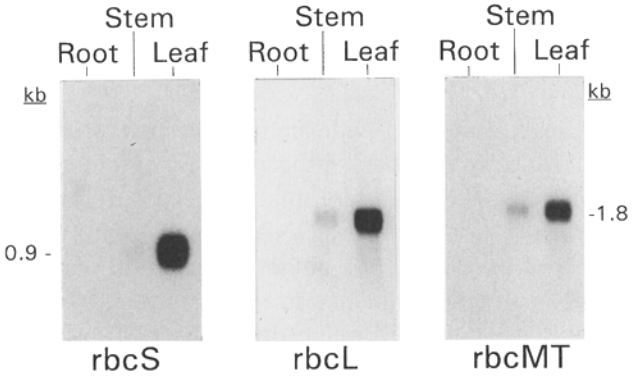


Fig. 4. Organ-specific accumulation of rbcMT mRNA. Messenger-RNA was isolated from roots, stems, and leaves of 10-day-old chamber-grown pea. Northern blots were loaded on an equal RNA basis and were probed with radiolabeled antisense RNA to *rbcS*, *rbcL* or rbcMT. Northern blots of *rbcS*, *rbcL* and rbcMT mRNA were exposed to X-ray film for 2 h, 1 h, and 36 h, respectively.

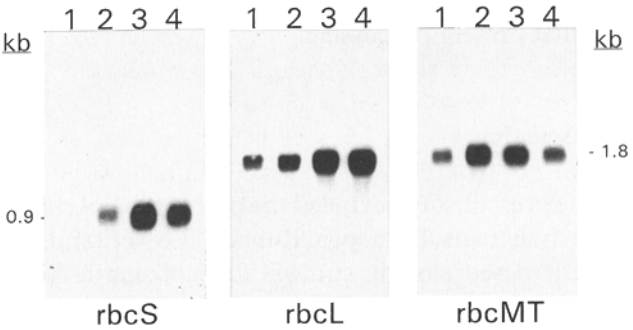


Fig. 5. Light-dependent accumulation of rbcMT mRNA in etiolated pea. Peas were germinated in a dark chamber in a light-tight room. After 8 days, etiolated seedlings were either harvested (treatment 1) or transferred to the light for 24 h (treatment 2) or 72 h (treatment 3). Control seedlings were germinated in the light and harvested after 8 day (treatment 4). RNA was isolated from leaf tissue from each treatment and northern analyses were conducted. Northern blots of *rbcS*, *rbcL*, and rbcMT were exposed to X-ray film for 1 h, 1 h, and 36 h, respectively.

Rubisco LSMT enzyme increased during the greening of dark-grown peas from 11 pmole CH_3 per minute per mg protein in dark-grown leaves to an apparent maximum of 32.5 pmole CH_3 per minute per mg protein after 72 h illumination. This level of extractable Rubisco LSMT enzyme activity was similar to that observed (32.4 pmole CH_3 per minute per mg protein) for peas grown for eight days under continuous illumination. Interestingly, the level of rbcMT mRNA in continuous illuminated leaves was significantly lower than the levels observed during the early stages of greening of pea (lanes 2 and 3 vs. 4). In fact, levels of rbcMT mRNA from continuous illuminated plants was not visibly different from dark-grown leaves. As expected, *rbcS* and *rbcL* transcript levels also increased upon illumination of dark-grown seedlings. In contrast to rbcMT, transcripts of *rbcS* and *rbcL* reached an apparent maximum during the latter stages of greening (lane 3). In addition, *rbcS* and *rbcL* transcript levels remained elevated in leaves grown under continuous illumination (lane 4). These results indicate that, unlike *rbcS* and *rbcL*, transcript levels for rbcMT reach an apparent maximum during the early stages of light-induced leaf development and decline in mature light-grown leaf tissue. These changes in transcript levels would be expected for an enzyme whose function involves post-translational protein processing.

Discussion

The present study detailed the purification of peptic fragments from pea Rubisco LSMT and a PCR-based cloning strategy for isolating a full-length cDNA. A similar strategy was previously utilized to obtain a full-length cDNA of sucrose-phosphate synthase from spinach [22]. The low abundance of Rubisco LSMT in pea leaves (ca. 0.01%) prompted the use of PCR, since it would be more difficult to obtain enough protein to ensure the production of an antibody with high titer and specificity with which to screen a library. Further, the protein sequence information obtained from peptic fragments permitted the confirmation of clones encoding for Rubisco LSMT. Hence, a molecular probe of the pea rbcMT gene was rapidly obtained thereby permitting identification of protein and nucleotide sequence, and characterization of rbcMT gene expression.

To date, the deduced amino-acid sequence of Rubisco LSMT represents the first reported example of a protein N-methyltransferase. Thus, it is now possible to extend the comparison of known enzyme sequences to include this class of methyltransferases. Interestingly, the deduced amino acid sequence of Rubisco LSMT does not possess any of the three sequence motifs proposed by Kagan and Clarke [20] for methyltransferases. However, knowledge of methyltransferase sequences is still fragmentary and no sequences are yet available for protein arginine, histidine, or N-terminal amino N-methyltransferases. As noted by Kagan and Clarke [20], methyltransferases whose sequences are available are limited and a number of other methyltransferases apparently do not possess the proposed motifs or any additional elements of sequence similarity. Furthermore several lines of evidence suggest that Rubisco LSMT exclusively methylates the large subunit of Rubisco [17, 18]. This high level of specificity may in part explain the lack of overall homology with other methyltransferases. Hence, sequence determination of other yet-to-be-discovered protein(lys)N-methyltransferases may be necessary to identify conserved, functionally essential regions in this class of enzyme.

Several lines of evidence indicate that there is a low copy number of the rbcMT gene in pea. Genomic Southern blot analysis revealed simple hybridization patterns. DNA sequence information of several cDNA clones revealed an invariant nucleotide sequence in the coding and non-coding regions. Although these observations do not preclude the existence of multiple structural genes encoding Rubisco LSMT, they are consistent with a low- or even single-copy gene hypothesis.

Many plant genes are expressed in a highly regulated manner. Gene products may be present only in certain cell types, at specific stages of

development or only following the application of distinct environmental stimuli [25]. In addition, the expression of nuclear genes encoding plastid proteins is often coordinated with the expression of plastid-encoded protein subunits [35]. The present study reveals that rbcMT gene expression is regulated in an organ-specific manner at the level of transcription or mRNA stability. The organ-specific expression of rbcMT paralleled that of *rbcS* and *rbcL* being predominately localized to photosynthetic leaf tissue. Examination of transcript levels during the light-induced development of etiolated pea leaves indicated that accumulation of mRNA encoding for *rbcS*, *rbcL*, and rbcMT is light-dependent. However, the activation of rbcMT expression preceded the maximum accumulation of mRNA encoding for either of the Rubisco subunits. Maximum transcript levels for rbcMT were obtained in the first 24 h of illumination which corresponded with the initial, light-dependent phase of *rbcS* and *rbcL* transcript accumulation. Interestingly, the kinetics of Rubisco activase mRNA accumulation during the greening of etiolated barley was similar to that reported here for rbcMT mRNA [47]. We also observed that in continuously illuminated peas leaves, rbcMT transcript levels were equal to the levels observed in dark-grown leaves (Fig. 5), while the activity of Rubisco LSMT was nearly 3-fold higher. Since the relative amounts of rbcMT transcripts increased dramatically during the initial phase of light-induced development of etiolated pea leaves, and then declined to a level equal to the those observed in the dark, changes in the level of Rubisco LSMT protein may be controlled by the level of rbcMT transcripts.

Finally, while a number of N-methylated lysyl residues in several proteins have been described, no unifying hypothesis with regards to the functional significance of methylated lysyl residues has been discovered. Molecular studies have approached this topic by engineering amino acid substitutions at the position of the methylated lysyl residues in calmodulin [36] and cytochrome *c* [3, 4], followed by expression of these mutant proteins in transformed tobacco plants and yeast cells, respectively. While the mutated calmodulin and cytochrome *c* proteins were incapable of acting as substrates for methylation, these studies were inconclusive as to a clear role for site-specific methylation of the target lysyl residues by the calmodulin or cytochrome *c* protein specific N-methyltransferases. With the cloning of the Rubisco LSMT gene from pea, more elegant and direct *in vitro* and *in vivo* analyses of the functional significance of methylation of Lys-14 in the LS of Rubisco can be conducted. Studies utilizing transgenic plants are presently underway to address the biochemical and physiological consequences of perturbing the methylation state of Rubisco without changing the amino acid sequence of the LS. Additionally, as protein and DNA sequences become available for other protein methylase III enzymes, the possibility that each of the protein N-methyltransferases coevolved with their respective protein substrates [16], thus explaining their exceptionally high degree of protein substrate specificity, can be addressed.

Acknowledgements

We thank Dr S.J. Gantt (University of Minnesota, Twin Cities Campus, Minneapolis, MN) for kindly providing a λgt10 pea leaf cDNA library. Special thanks to M. Royer and C. Patterson for technical assistance. This work was supported by US Department of Energy Grant DE–F05–92ER20075 to R.L.H. The investigation reported in this paper (94–10–164) is in connection with a project of the Kentucky Agriculture Experiment Station.

References

1. Andrews TJ, Lorimer GH: Rubisco: structure, mechanisms, and prospects for improvement. In: Hatch MD, Boardman NK (eds) The Biochemistry of Plants, vol. 10, pp. 131–218. Academic Press, New York (1987).
2. Bedbrook J: A plant nuclear DNA preparation procedure. Plant Mol Biol Newslett 2: 24 (1981).
3. Ceesay KJ, Rider LR, Bergman LW, Tuck MT: The relationship between the trimethylation of lysine 77 and

cytochrome *c* metabolism in *Saccharomyces cerevisiae*. Int J Biochem 26: 721–734 (1994).
4. Cessay KJ, Bergman LW, Tuck MT: Further investigations regarding the role of trimethyllysine for cytochrome *c* uptake into mitochondria. Int J Biochem 23: 761–768 (1991).
5. Chang FN, Cohen LB, Navickas IJ, Chang CN: Purification and properties of a ribosomal protein methylase from *Escherichia coli* Q13. Biochemistry 14: 4994–4998 (1975).
6. Clarke S: Protein carboxyl methyltransferases: two distinct classes of enzymes. Ann Rev Biochem 54: 479–506 (1985).
7. Dean C, Pichersky E, Dunsmuir P: Structure, evolution, and regulation of *rbcS* genes in higher plants. Annu Rev Plant Physiol Plant Mol Biol 40: 415–439 (1989).
8. DiMaria P, Kim S, Paik WK: Cytochrome *c* specific methylase from wheat germ. Biochemistry 21: 1036–1044 (1982).
9. Farooqui J, DiMaria P, Kim S, Paik WK: Effect of methylation on the stability of cytochrome *c* of *Saccharomyces cerevisiae* in vivo. J Biol Chem 256: 5041–5045 (1981).
10. Farooqui J, Kim S, Paik WK: In vivo studies on yeast cytochrome *c* methylation in relation to protein synthesis. J Biol Chem 255: 4468–4473 (1980).
11. Frost B, Paik WK: Cytochrome *c* methylation. In: Paik WK, Kim S (eds) Protein Methylation, pp. 59–76. CRC Press, Boco Raton, FL (1990).
12. Frost B, Syed SK, Kim S, Paik WK: Effect of enzymatic methylation of cytochrome *c* on its function and synthesis. Int J Biochem 22: 1069–1074 (1990).
13. Gantt JS, Baldauf SL, Calie PJ, Weeden NF, Palmer JD: Transfer of *rpl22* to the nucleus greatly preceded its loss from the chloroplast and involved the gain of an intron. EMBO J 10: 3073–3078 (1991).
14. Gregori L, Marriott D, Putkey JA, Means AR, Chau V: Bacterially synthesized vertebrate calmodulin is a specific substrate for ubiquitination. J Biol Chem 262: 2562–2567 (1987).
15. Gregori L, Marriott D, West CM, Chau V: Specific recognition of calmodulin from *Dictyostelium discoideum* by the ATP ubiquitin-dependent degradative pathway. J Biol Chem 260: 5232–5235 (1985).
16. Han C-H, Richardson J, Oh S-H, Roberts DM: Isolation and kinetic characterization of the calmodulin methyltransferase from sheep brain. Biochemistry 32: 13974–13980 (1993).
17. Houtz RL, Poneleit L, Jones SB, Royer M, Stults JT: Posttranslational modifications in the amino-terminal region of the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from several plant species. Plant Physiol 98: 1170–1174 (1992).
18. Houtz RL, Royer M, Salvucci ME: Partial purification and characterization of ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit $^{\epsilon}$N-methyltransferase. Plant Physiol 97: 913–920 (1991).
19. Houtz RL, Stults JT, Mulligan RM, Tolbert NE: Post-translational modifications in the large subunit of ribulose bisphosphate carboxylase/oxygenase. Proc Natl Acad Sci USA 86: 1855–1859 (1989).
20. Kagan RM, Clarke S: Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. Arch Biochem Biophy 310: 417–427 (1994).
21. Keegstra K, Olsen LJ: Chloroplastic precursors and their transport across the envelope membranes. Annu Rev Plant Physiol Plant Mol Biol 40: 471–501 (1989).
22. Klein RR, Crafts-Brandner SJ, Salvucci ME: Cloning and developmental expression of the sucrose-phosphate-synthase gene from spinach. Planta 190: 498–510 (1993).
23. Klein RR, Mullet JE: Light-induced transcription of chloroplast genes. J Biol Chem 265: 1895–1902 (1990).
24. Klein RR, Salvucci ME: Photoaffinity labeling of mature and precursor forms of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase after expression in *Escherichia coli*. Plant Physiol 98: 546–553 (1992).
25. Kuhlemeier C, Green PJ, Chua N-H: Regulation of gene expression in higher plants. Annu Rev Plant Physiol 38: 221–257 (1987).
26. Lobet Y, Lhoest J, Colson C: Partial purification and characterization of the specific protein-lysine N-methyltransferase of YL32, a yeast ribosomal protein. Biochim Biophys Acta 997: 224–231 (1989).
27. McCurry SD, Gee R, Tolbert NE: Ribulose-1,5-bisphosphate carboxylase/oxygenase from spinach, tomato or tobacco leaves. Meth Enzymol 90: 515–521 (1982).
28. Mudgett MB, Clarke S: Characterization of plant *L*-isoaspartyl methyltransferases that may be involved in seed survival: purification, cloning, and sequence analysis of the wheat germ enzyme. Biochemistry 32: 11100–11111 (1993).
29. Mullet JE: Chloroplast development and gene expression. Annu Rev Plant Physiol Plant Mol Biol 39: 475–502 (1988).
30. Mulligan RM, Houtz RL, Tolbert NE: Reaction-intermediate analogue binding by ribulose bisphosphate carboxylase/oxygenase causes specific changes in proteolytic sensitivity: the amino-terminal residue of the large subunit is acetylated proline. Proc Natl Acad Sci USA 85: 1513–1517 (1988).
31. Mutoh N, Simon MI: Nucleotide sequence corresponding to five chemotaxis genes in *Escherichia coli*. J Bact 165: 161–166 (1986).
32. Paik WK, Nochumson S, Kim S: Carnitine biosynthesis via protein methylation. Trendo Biochem Sci 2: 159–162 (1977).
33. Paik WK, Kim S: Protein methylation. In: Freedman KB, Hawkins HC (eds) The Enzymology of Posttranslational Modifications of Proteins, vol. 2, pp. 187–228. Academic Press, London (1985).
34. Patterson SD, Hess D, Yungwirth T, Aebersold R: High-

yield recovery of electroblotted proteins and cleavage fragments from a cationic polyvinylidene fluoride-based membrane. Anal Biochem 202: 193–203 (1992).
35. Rapp JC, Mullet JE: Chloroplast transcription is required to express the nuclear genes *rbcS* and *cab*. Plant Mol Biol 17: 813–823 (1991).
36. Roberts DM, Besl L, Oh S-H, Masterson RV, Schell J, Stacey G: Expression of a calmodulin methylation mutant affects the growth and development of transgenic tobacco plants. Proc Natl Acad Sci USA 89: 8394–8398 (1992).
37. Roberts DM, Rowe PM, Siegel FL, Lukas TJ, Watterson DM: Trimethyllysine and protein function. J Biol Chem 261: 1491–1494 (1986).
38. Rowe PM, Wright LS, Siegel FL: Calmodulin N-methyltransferase. J Biol Chem 261: 7060–7069 (1986).
39. Sambrook J, Fritsch EF, Maniatis T: Molecular cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
40. Sarnow P, Rasched I, Knippers R: A histone H4-specific methyltransferase properties, specificity and effects on nucleosomal histones. Biochim Biophys Acta 655: 349–358 (1981).
41. Theg SM, Scott SV: Protein import into chloroplasts. Trends Cell Biol 3: 186–190 (1993).
42. Tuck MT, Farooqui JZ, Paik WK: Two histone H1-specific protein-lysine N-methyltransferases from *Euglena gracilis*. J Biol Chem 260: 7114–7121 (1985).
43. von Heijne G, Nishikawa K: Chloroplast transit peptides: the perfect random coil? FEBS Lett 278: 1–3 (1991).
44. Wang P, Royer M, Houtz RL: A novel affinity purification technique for ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit $^{\epsilon}$N-methyltransferase. Prot Expr Pur submitted (1994).
45. Ziegenhagen R, Goldberg M, Rakutt WD, Jennissen HP: Multiple ubiquitination of calmodulin results in one polyubiquitin chain linked to calmodulin. FEBS Lett 271: 71–75 (1990).
46. Ziegenhagen R, Jennissen HP: Plant and fungus calmodulins are polyubiquitinated at a single site in a Ca^{2+}-dependent manner. FEBS Lett 273: 253–256 (1990).
47. Zielinski RE, Werneke JM, Jenkins ME: Coordinate expression of rubisco activase and rubisco during barley leaf cell development. Plant Physiol 90: 516–521 (1989).